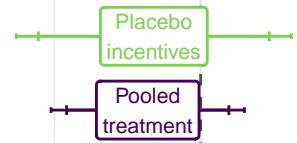
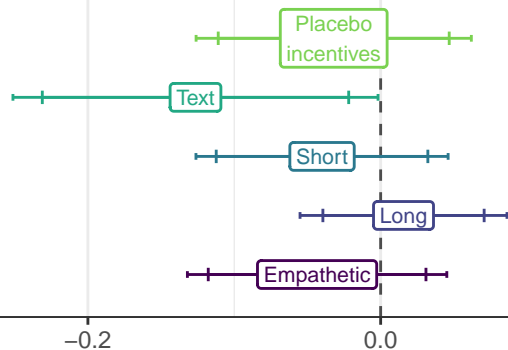


Pooled



Disaggregated



$p(\text{Short} > \text{Text}) = 0.10$
 $p(\text{Long} > \text{Text}) = 0.01$
 $p(\text{Empathetic} > \text{Text}) = 0.11$
 $p(\text{Empathetic} > \text{Short}) = 0.53$
 $p(\text{Empathetic} > \text{Long}) = 0.89$
 $p(\text{Long} = \text{Short}) = 0.23$